DOCUMENT RESUME

ED 347 990 IR 015 727

AUTHOR Gleim, Candice W.; Harvey, Francis A.
TITLE Viewing Hypermedia through the Prism of Evaluation.
PUB DATE Feb 92
NOTE 18p.; In: Proceedings of Selected Research and Development Presentations at the Convention of the Association for Educational Communications and Technology and Sponsored by the Research and Theory Division; see IR 015 706.
PUB TYPE Information Analyses (070) -- Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Computer Assisted Instruction; Elementary Secondary Education; Engineering Education; *Evaluation Methods; *Hypermedia; Instructional Effectiveness; Meta Analysis; Postsecondary Education; *Research and Development; *Research Methodology; Technology Transfer
IDENTIFIERS *Lehigh University PA

ABSTRACT

This paper seeks to establish evaluation as an important concern in hypermedia research and development, and presents a review of current hypermedia evaluation research with a prescriptive focus. Characteristics of hypermedia currently attracting the attention of instructional developers are discussed and conclusions drawn from recent evaluation research are presented along with distinguishing features that separate the evaluation approaches employed. Characteristics of hypermedia that present challenges to evaluative studies are analyzed and possible methodological solutions are suggested. Current and proposed activities at Lehigh University (Pennsylvania) which seek to apply the ideas presented in this paper to the development of hypermedia applications for technology transfer and training in engineering and other areas are also described. It is concluded that hypermedia requires an adaptive interactive research model to investigate the effects of treatment, and that this model will probably incorporate elements of both rationalistic and naturalistic approaches. Two figures show a comparison of rationalistic and naturalistic methodologies and the relationship between learner control and self-regulation. Three tables display the characteristics of hypermedia and expected outcome; selected evaluation research studies; and criteria for judging the trustworthiness of the findings of an evaluation study. (50 references) (BBM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED347990

Title:

Viewing Hypermedia Through the Prism of Evaluation

Authors:

Candice W. Gleim
Francis A. Harvey

IR015727

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Michael Simonson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

Viewing Hypermedia Through the Prism of Evaluation

Candice W. Gleim

Francis A. Harvey

Lehigh University

Introduction

Hypertext, a vision conceived in the interest of information management (Bush, 1945), has had an influential effect on the "look and feel" of computer interfaces, creating a fusion of once isolated applications and information sources. Hypertext (or more generally hypermedia when encompassing video, sound, and graphics as well as text) not only supports the organization and management of information, but also facilitates the presentation and exploration of information. It is the latter capability that has stimulated considerable debate among educators concerned with the theory and practice of technology-based instruction. The user-centered experiential nature of hypermedia technology makes it a likely vehicle for delivery of instruction (Allison & Hammond, 1990). Unfortunately, the development of hypermedia-based instructional delivery systems is hampered by a scarcity of research which can be used to guide practice and with problems associated with the open-ended design.

Evaluation of technology-based instructional programs often provides insightful investigation of an innovation concerning salient features of the technology and how the features interact in an applied setting. However, evaluation research is inherently problematic due to competing interests of organizations and individuals involved in the process; these problems are not discussed in this paper (see Rossi & Freeman, 1989; Reeves, 1991). Instead, the purpose of this paper is twofold: we seek to establish evaluation as an important concern in hypermedia research and development, and to present a review of current hypermedia evaluation research with a prescriptive focus. Characteristics of hypermedia currently attracting the attention of instructional developers will be discussed. Unfortunately when the attributes of hypermedia are brought into focus by recent evaluation research, findings are often limited; conclusions drawn from these developments will be presented, along with distinguishing features that separate the evaluation approaches employed. Characteristics of hypermedia which present challenges to evaluative studies will be analyzed and possible methodological solutions will be raised. Finally, we will describe current and proposed activities at Lehigh which seek to apply the ideas presented in this paper to the development of hypermedia applications for technology transfer and training in engineering and other areas.

Hypermedia and Computer-Assisted Learning

The interface can be thought of as the first gateway to effective instructional delivery in a computer-assisted learning (CAL) system. In contrast to earlier behavioristic models which required the learner to respond in a specific manner and measured observed behavior (Skinner, 1986), more recent delivery systems influenced by a cognitive approach attempt to provide learning experiences for the learner so that the learner may elaborate, modify and integrate the material presented with previous knowledge. The experiential nature of hypermedia meshes with the exploratory and generative learning orientation of cognitive theory. Consequently hypermedia is becoming an attractive choice for some CAL developers (Allison & Hammond, 1990). However, the open environment of hypermedia presents new challenges in designing and evaluating applications specifically to achieve learning goals.

The principal problem related to using hypermedia as a delivery mechanism for instruction, as identified by Hammond (1989), concerns the learner's experience exploring the information. Exposure to large bodies of unfamiliar information may not be suitable for all tasks or all individuals and may result in confusion or indecisiveness on the part of the learner. Developing an overview of the material and the way it is organized can be difficult for the learner; sometimes merely locating information that the individual knows is included in the application can be a problem. Connecting the information in a myriad of ways may create uncertainty for learners, who may become hesitant to make a selection because they are unsure of what to expect or of what

is expected of them. Cognitive resources expended in attempting to master the hypermedia interface may interfere with the intended learning task.

Characteristics of Hypermedia Applications Designed for Learning

Currently there are no universal standards governing what hypermedia (or hypertext) entails, although efforts in this direction are being formulated (Moline, Benigni, & Baronas, 1990). The instructional developer must depend on experience and intuition to design and develop an application incorporating features of hypermedia or hypertext. Therefore, characteristics of hypermedia for learning as identified in literature on the use of hypermedia in learning environments will be explored as potential variables of evaluation studies.

Gagné and Briggs (1979) identify four classes of variables that have the potential to influence an evaluation study: process variables, outcome variables, support variables, and aptitude variables. Process variables refer to factors in an instructional situation which have the potential to influence outcome. Outcome variables refer to measured variables; the focus of the study. Support variables are represented as "opportunities for learning" (p. 297); for example, the presence of adequate materials and climate represent such variables. Aptitude variables represent the learner's ability in an educational situation.

Process variables include: instructional sequence, pacing, and events of instruction (Gagné & Briggs, 1979). Romiszowski (1990) suggests the focus of hypermedia literature emphasizes the benefits of hypermedia for learning (e.g. non-linear access, learner-controlled, distributed organization) resulting in a "process oriented philosophy" (p. 322). Viewing hypermedia in a process oriented context, as characterized in the literature, separates each characteristic from the expected outcome or goal anticipated by using hypermedia for learning.

The products of learning, or outcome variables, are often characterized in cognitive terms in the literature describing the benefits of hypermedia. For example, Kozma (1991) suggests that knowledge representation may exist along multiple instead of single conceptual dimensions. He also suggests that complex situations generate an assembly of representations instead of retrieving an intact schema; hypermedia may facilitate this representation. However, traditional tests are not designed to provide such cognitive assessments.

Attributes which emphasize the benefit of using hypermedia for learning as characterized in literature are identified in Table 1 as primary process characteristics. Each primary process characteristic (i.e., independent of other characteristics) is presented with an associated product characteristic. Secondary characteristics are not identified in this table and include those factors that are influenced by and have the potential to influence additional characteristics. An example of a secondary characteristic is self-regulation; Kinzie and Berdel (1990) suggest self-regulation is facilitated by learner control. To elucidate this distinction, an example representing the relationship between learner-control and self-regulation as discussed in Kinzie and Berdel is represented in Figure 2. Although Kinzie and Berdel suggest hypermedia may facilitate self-regulation, it is mediated through learner control and therefore does not appear as a characteristic in Table 1. A perspective of these characteristics is available through the filter evaluation provides.

Recent Developments in Evaluation Research on Hypermedia

Hypermedia, as it pertains to instructional systems technology, incorporates many aspects associated with computer-based instruction (CBI). Research conducted on factors such as color, screen arrangement, video, and learner control, are applicable to hypermedia design. Evaluation of hypermedia applications becomes more problematic than directed CBI because the major characteristic of hypermedia, the ability for the learner to traverse the application in a multitude of ways, makes each learner's experience unique. Browsing and navigation, inherent in hypermedia, decrease the ability to control the experimental condition.

Our objective is not to present a review of all hypermedia and hypermedia-related evaluative research. Instead, we wish to concentrate on studies of hypermedia evaluative research used in educational settings. A second criterion, educational applications as opposed to training applications, further narrowed the literature selection process. Table 2, and 2a summarize the following information for each of the eleven articles (consisting of 15 studies) identified by the criteria: evaluation approach used, focus of the study, indication pertaining to support of research hypothesis, instructional content, subjects, and an indication representing the number of treatments.

Conclusions Drawn from Evaluation Research on Hypermedia

Results from hypermedia evaluation studies, identified by selection criteria, are consolidated into five conclusions. Each conclusion is accompanied with an associated cautionary note. Conclusions and cautionary notes are discussed below.

Conclusion 1. Hypermedia study guides can be successfully used as a supplement to and substitute for lecture instruction (Higgins & Boone, 1990; Horton, Boone, & Lovitt, 1990). The consistent delivery inherent in computerized instruction, offers remediation for students needing additional instruction to master subject content, or assistance to students that have missed a lesson. Learner-initiated control of content through a hypermedia interface provides self-pacing, which may be particularly helpful for learning disabled students. A range of instructional cues (in the form of directed assistance) depending on responses to multiple-choice questions, match instructional remediation level with learner's highest level of independent comprehension.

Caution 1. Objective outcome measurements used in comparative studies (Higgins & Boone, 1990), require instructional overlap between the computer-assisted content and supplementary lesson materials such as textbook or teacher initiated lecture. Scarcity of commercially available software or resources necessary for development of quality instruction which match specific content needs of the instructor are prohibitive factors. Measured gain for each learner may differ depending on: learner educational ability, relationships between various screens, and granularity of the nodes in the hypermedia instruction (cf. Lanza & Roselli, 1991; Tripp & Roby, 1990).

Conclusion 2. Learning performance skills (e.g. problem solving, management practices, group interaction) which require a vast range of experiences can be mediated through episodes of peer modeling using video illustrations (Goldman & Barron, 1990). Lesson planning time for the instructor is easier when supplementary materials exist on videodisc.

Caution 2. Development and equipment costs for hypermedia may far exceed costs of traditional instructional methods. A "bottleneck," from new instructional remedy to technological-based teaching tool, may discourage experimenting with new solutions to problems. Although student performance skills increased, students were apprehensive about preparing for traditional examinations after hypermedia instruction was substituted for the lecture (Goldman & Barron, 1990). Students may feel more confident about the importance of information when it is traditionally delivered.

Conclusion 3. Evaluation of hypermedia enhancements during the development cycle has the potential to identify deficiencies of the product, or foster new ideas to improve learne rategies during the development stage (Egan et al., 1989; Covey, 1990). Computer-based monitoring of the frequency with which the learner uses available strategies and of learner task success rates provide a rich source of information to guide subsequent development. Affective measures obtained early in the development process may eliminate costly ineffective options (e.g. such as digital audio feedback, touch screens, etc.) based on learner reactions to pilot versions.

Caution 3. Altering more than one function in an instructional application between formative evaluations may confound the ability to identify the source of improvement. Studies of this nature require a homogeneous pool of subjects. Although subjects may be comparable with respect to many demographic variables, groups of subjects may differ in aspects not compared (e.g. computer experience, motivation, anxiety, history).

Conclusion 4. Metaphors provide a meaningful way to integrate a wide range of information connected by theme (Gay, Trumbull, & Mazur, 1991). The learner is able to utilize the information in an intuitive way.

Caution 4. Post treatment interviews can provide insightful affective indicators pertaining to techniques used in the application, but matching other characteristics of the learner with comments that arise during the interview (e.g. prior knowledge of the subject matter, ability to direct learning, and problem solving skills) makes generalization difficult. The use of metaphor has not been successful in all systems (Tripp & Roby, 1990). When combining metaphors and graphical browsers (or using browsers alone, Mays, Kibby, & Watson, 1988), elements may compete for learners' attention.

Conclusion 5. Relating a series of phenomena encourages interrelational thinking (Landow, 1989). As demonstrated through class discussions and examinations, students contributed a greater range of information than previous classes. Examination responses were judged to be more intellectually sophisticated and more rigorous than responses from previous

classes. Students appeared more motivated to contribute their ideas of inter-connectivity, and were perhaps more apt to complete assigned readings because they felt all assignments were interrelated.

Caution 5. Longitudinal studies may indicate novelty or duration limitations. Implementing a study which evaluates performance over a period of treatments may provide evidence that performance using hypermedia increases over time (Higgins & Boone, 1990). Direct transfer of hypermedia experience to course content may occur more frequently when content included in hypermedia instruction is contributed based on the recommendations of the instructor. Conceptual arrangement of the hypermedia, selected by the instructor or designer, may have an effect on learner's ability to make connections. Giving learners the ability to create their own links may influence positive reaction to hypermedia. As *Caution 1* mentioned, developmental costs and instructor's expertise may limit direct transfer of any computer-based instruction to course content.

These conclusions, as shown in Table 2, Table 2a, encompass different approaches. The next section compares different evaluation approaches used by researchers as outlined in the tables.

Evaluation Research Methodologies Characterized and Discussed

Several studies did not utilize an objective outcome variable to measure learning (Landow, 1989; Gay, Trumbull, & Mazur, 1991; Egan et al., 1989; Covey, 1990). Instead, the study focused on outcome which measured the success of searching through a hypermedia application (Gay, Trumbull, & Mazur, 1991; Egan et al., 1989), or used an essay exam which was judged with interrater reliability measures (Landow, 1989; Egan et al., 1989). With such a rich open-ended learning environment, what approach should be used to measure the impact of the experience on the learner?

The methodological question is often answered by inquirer, stakeholder, board members, community members, or the editor responsible for judging the inquiry (Guba & Lincoln, 1984). Often professionals will select an approach practiced within a tradition (cf. Jacob, 1987; Atkinson, Delamont, & Hammersley, 1988); this has practical value as researchers within a professional community can exchange results. Each approach has its advantages, but two general approaches as mentioned by Guba and Lincoln (1984), rationalistic[1] and naturalistic, in many ways can be characterized as dichotomies (see Figure 1). Although they are often mentioned as discrete categories, they may be more appropriately considered as ends of a continuum.

Regardless of the form of inquiry, Guba and Lincoln (1984) suggest that each approach has a set of criteria broadly defined into four categories which can be used to identify the trustworthiness of the study's findings (see Table 3). Truth value represents the confidence of the findings; applicability describes suitability of findings to similar inquiries with different setting/subject arrangement; consistency refers to the ability to replicate the study; and neutrality represents the degree to which the findings are established without biases.

Classified as rationalistic methods by Guba and Lincoln (1984), **experimental and quasi-experimental inquires** have a long tradition modeled after scientific forms of inquiry (Campbell & Stanley, 1963; Cook & Campbell, 1979). The distinction between the two rests on scheduling, either measurement or exposure. Where experimental designs are more tightly controlled, experimental rigor is not abandon in situations where, for example, random assignment to treatment is impossible. Instead quasi-experimental techniques are employed.

Formative evaluation was used in a few of the evaluation studies reviewed. Labeled a proactive approach by Stufflebean (1984), this technique offers systematic indication of the merits of specific options prior to completion of product development, in contrast to summative design which evaluates at the end of development.

Although Stufflebean's (1984) and Scriven's models go beyond mere product evaluation, Dick and Carey (1990) prescribe formative approaches that are appropriate for instructional products. The formative evaluation draws upon the systematic practices of quasi-experimental and experimental design. Yet, in practice, it takes on a range of techniques (cf. Covey, 1990; Egan et al., 1989). This technique seeks convincing evidence (Gagné & Briggs, 1979) and hence lacks correlates for all categories for comparison in Table 3.

Naturalistic inquiries use several techniques to guard against accusations of less rigor than rationalistic studies. For example, triangulation, a mechanism to cross-verify data with professionals of varying perspectives, is analogous to objectivity in rationalistic studies (Guba & Lincoln, 1984).

BEST COPY AVAILABLE

Interrelationship of Characteristics of Hypermedia and Results from Evaluation Studies

Landow (1989) found, through naturalistic inquiry, an improvement in the quality and depth of responses to essay exams. Providing the learner with extension and construction ability (Kozma, 1991) may increase the learner's capacity to interrelate ideas. Node size and representation provided by links may also contribute to success. However, currently there is no evidence to support the notion that experience with hypermedia fosters transfer of learners' capacity to interrelate ideas in new situations.

As suggested by Landow (1989), hypermedia does facilitate learners' ability to control their own learning (Marchionini, 1988;Kinzie & Berdel, 1990; Hammond, 1989), but studies by Higgins and Boone (1990) found differences in the rate of gain between groups of different ability. Is the difference due to the interacting effects of: organization of the hypermedia application, control strategies utilized, and aptitude level of the learner? Is this control best fostered when it is used to control the content, or when it is used to control instructional strategies (e.g. examples, practice questions, amount of time needed, Merrill, 1983)?

Experimental and quasi-experimental approaches provide the researcher with a tradition of experience with which to draw upon. However, they have failed education with respect to comparison studies, an approach that is often used to investigate the effects of technology and traditional forms of teaching (Reeves, 1991;Clark, 1985; Clark, 1983). Using traditional approaches to design research studies require studies to be designed, planned, and implemented to maximize experimental variance (Kerlinger, 1973). Traditional approaches to evaluation may never provide a clear picture of instructional innovations when the focus of the study concerns process characteristics that require comparison studies to investigate the effects of the treatment (e.g. linear over nonlinear acquisition, cf. Lanza & Rozelli, 1991; Mayes, Kibby & Watson, 1988). The quasi-experimental approach may provide effective research designs when research questions require temporal probes. Single group designs exist to measure persistence (or gain) of each treatment with a control measure to monitor maturity (e.g. design 8, Campbell & Stanley, 1963), or when an indication of comparative effects of distinctively different treatments is desired (Atkinson, 1968; Borg, 1984).

The disadvantage of quasi-experimental evaluation rests on the additional attention the evaluator must place on the threats to internal validity (without randomizing subjects to study group), or external validity (time requirement to conduct such a study), and the statistical requirement of reliable measures of outcome.

If hypermedia is viewed as a cognitive enhancer, must we change our instructional objectives from performance-oriented to content-oriented as Bonner (1988) suggests? Self-tests, consisting of synthesizers or analyzers, may aid self-discovery while tracing the learner's progress, providing developers with information that may assess progress. Gay and Raffensperger (1989) mention a creative alternative to objective measures giving students the capability to self-test their progress. It consists of a video segment with an instructional overlay. The learner is invited to self-test understanding in a graphic environment by manipulating variables, thus demonstrating understanding of the instruction.

A cognitive approach to instruction may reestablish when instructional prescriptions are defined (Winn, 1990). Winn asserts that prescriptions for cognitive instructional strategies are generated while the instruction is underway, as opposed to defining them early in the development cycle, as in behavioral instructional strategies. The teacher or designer must appropriately time the necessary intervention when practicing a cognitive approach to instruction. Either must monitor the engaged learner and provide mechanisms to change strategies when appropriate. This also suggests a formative design strategy to insure that the designer can monitor actual use of the system after implementation and respond accordingly with necessary changes based on actual use.

Planned interventions, depending on the way they are implemented (e.g., by the teacher instead of the application) may alter learners' success depending on attributes present at the time of the intervention. Evaluating additive effects of the interaction may require a model such as those used in library science (Saracevic, Mokros, & Su, 1990). This form of general model-specific model captures and measures the attributes of the interaction process and specific attributes of members participating in the interaction. This naturalistic approach, provides an analysis of what transpired during interaction without loss of descriptive power which typically occurs in a statistical study.

The effect of hypermedia on learner strategy formulation, although posed by many researchers when hypothesizing the value of hypermedia as an instructional medium, has not been evaluated. Similarly, evaluation of levels of instructional support is needed. Although hypermedia is designed to inspire the learner to control his/her own learning, instructional strategies designed to provide limited guidance, yet unending support, may remedy the problems associated with hypermedia.

Evaluation of Hypermedia at Lehigh University

Since September, 1989, faculty and staff of the Advanced Information Technologies Laboratory (AITL) of the College of Education at Lehigh University have been engaged in developing integrated hypermedia materials for a number of clients. The current major projects of the AITL, under the direction of Dr. Francis A. Harvey, are the Hypermedia Support System for the Bridge Fatigue Investigator (Hyper-BFI) and the Cities in Schools Hypermedia Development Project (CIS/HDP). In this section of the paper we will described current and planned evaluation efforts related to those projects.

Evaluation of the Hypermedia Support System for the Bridge Fatigue Investigator

The Hypermedia Support System for the Bridge Fatigue Investigator Project is a subproject of Lehigh's Center for Advanced Technologies for Large Structural Systems (ATLSS) sponsored by the National Science Foundation. The project is part of a program which, for the past five years, has been exploring the use of combinations of knowledge-based system technology and hypermedia technology for information utilization and dissemination, training, and technology transfer (see Harvey, 1990; Harvey, 1989).

The Bridge Fatigue Investigator (BFI) is an expert system designed to aid bridge inspection (Michalerya, Wilson, & Chen, 1990). An interdisciplinary team of subject matter experts (including the developers of BFI) and instructional designers is developing a hypermedia support system to train bridge inspectors in the use of BFI and to help professional engineers use BFI for post-inspection bridge analysis.

The Hyper-BFI Project is now in its beginning stages. Techniques used to support formative evaluation include using rapid prototyping to create a pre-prototype for team members to react to small slices of system development during team meetings. A special icon represented by an exclamation symbol is operational from any instructional screen. This is used to communicate problems with the screen directives, or navigational aids at any time during the course of instruction while a development team member is viewing the application. Additionally, a selected-response measure with Likert scale format is also being considered to gather feedback from users during pilot testing.

Evaluation design objectives at this stage concern local aspects of the system: the design of a reliable instrument which will measure learner attributions and beliefs with respect to the developing system (Clark, 1983), and the identification of specification misconceptions in design implementation. These evaluation mechanisms, where possible, will be incorporated into the pre-prototype hypermedia application as it is developed.

Evaluation of the Cities in Schools Hypermedia Development Project

The Cities in Schools Hypermedia Development Project has, for the past three years, been developing an integrated hypermedia training curriculum for Cities in Schools, Inc. (CIS), the nation's largest and most successful program addressing the needs of at-risk youth. A four-week, full-day integrated hypermedia training course for directors of CIS local projects has been developed which includes nearly 60 minutes of professionally produced motion video and approximately 500 still photographs on videodisc; over 500 pages of printed instructional resources and background material; and a comprehensive hypermedia computer program with several hundred computer-based instructional resource materials and navigational aids which allow both trainer and trainees to access, use, and interrelate computer-based, video, and print instructional materials (see Harvey, 1990; Harvey & Story, 1989).

A formative evaluation of the Cities in Schools Project Operations integrated hypermedia training course was carried out both during the training program for Cities in Schools trainers conducted at Lehigh in February, 1991 and during the initial training program delivered by Cities in Schools trainers in March, 1991 (Spokane, 1991). This study used both trained observers and

BEST COPY AVAILABLE

videotape to gather a comprehensive picture of how the training course was used and how trainers and trainees responded to it. These observations were supplemented by open-ended feedback forms completed by trainers and trainees both during and after training.

Since Cities in Schools training programs are held monthly in the National Center for Partnership Development hypermedia training laboratory at Lehigh, ample opportunity is provided to conduct both formative evaluation of CIS hypermedia materials under development and summative evaluation of those materials already in use. A follow-up evaluation of the Cities in Schools Project Operations training course is planned for March, 1992 at Lehigh.

The Cities in Schools integrated hypermedia training course for executive directors of city-wide CIS programs now under development will incorporate in its software unobtrusive mechanisms for recording complete information about how individual users make use of the hypermedia (Story & Harvey, 1990). Each interaction with the software (selection of icon, text entry, use of navigational support tools, etc.) will be recorded and time-coded automatically. A printed record of the interaction between user and the hypermedia program will be recorded on disk and later printed out for analysis.

In both of these projects, the evaluation components are designed at the same time as the application is designed. Design, development, and evaluation are seen as parallel and mutually supportive activities. This approach may ameliorate the "trapped administrator" syndrome identified by Campbell (1977), whereby the administrator is faced with the design of an evaluation study after the program is fully developed and in place.

Conclusion

Evaluation often distorts our view of instructional programs and the benefits provided by new technologies (Reeves, 1991). Consequently, developers of instructional software and educators alike use public receptivity as a barometer of success for their new products. Unfortunately, this offers minimal support in acknowledging the product as providing efficacy in instructional environments.

The benefit associated with hypermedia evaluation research lies in the dissemination of relevant findings and impact of the findings to enhance future practice. Research designs implemented with a rationalistic approach need valid constructs and measures to accurately quantify the variables related to the focus of the study. Plausible rival hypotheses can offer rich descriptions by contradicting a study's findings as a natural consequence of using rationalistic methods (Reeves, 1991).

Naturalistic approaches have the potential to contribute rich descriptions needed to foster new ideas and provide insight. Guba and Lincoln (1984) offer three tests to determine if a naturalistic approach should be selected over rationalistic methods: 1) data from study to study can not be aggregated; 2) researcher has little control of extraneous variance; and 3) impact of research on practice is absent. By these standards hypermedia, and many other focused studies of technology, is an ideal topic for naturalistic evaluation.

Despite the correlation of judgement criteria between rationalistic and naturalistic approaches (Table 3), the procedures are very different. Naturalistic study requires exceedingly large quantities of data as the approach is inductive instead of deductive. (In fact, hypermedia has been used to track and to interrelate the various data repositories in a naturalistic evaluation, see Martin, 1991.) Practitioners must become aware of naturalistic approaches to evaluation, their limitations and potential benefit to retrofit development. In particular, how would naturalistic inquiries remedy Guba and Lincoln's three-test scenario? Could data be aggregated for comparative analysis? Could researcher methods be analyzed? How would naturalistic studies influence practice?

The review and analysis in this paper suggests that hypermedia requires an adaptive interactive research model to investigate the effects of treatment. It is probable that this model will incorporate elements of rationalistic and naturalistic approaches. Regardless of the method used to evaluate the effectiveness of hypermedia, a focus on communicating the results of the study will foster progress in hypermedia development for learning.

References

Allison, Lesley, & Hammond, Nick (1990). Learning Support Environments: Rational and Evaluation. Computers in Education, 15, 137-143.

Atkinson, Paul, Delamont, Sara, & Hammersley, Martyn (1988). Qualitative Research Traditions: A British Response to Jacob, Review of Educational Research, 58,2, 231-250.

Atkinson, Richard C. (1968). Computerized Instruction and the Learning Process, American Psychologist, 23, 225-239.

Bonner, Jodi. (1988). Implications of Cognitive Theory for Instructional Design: Revisited. Educational Communications and Technology Journal, 36, 3-14.

Borg, Walter (1984). Dealing with Threats to Internal Validity That Randomization Does not Rule Out. Educational Researcher, December, 11-14.

Bush, Vannevar (1945). As We May Think. Atlantic Monthly 176 (1). 101-108.

Campbell, Donald T., & Stanley, Julian C. (1963). Experimental and Quasi-Experimental Designs for Research. Dallas: Houghton Mifflin Company.

Campbell, Donald T. (1977). Reforms as Experiments, Readings in Evaluation Research (2nd Ed.), NY: Russel Sage Foundation.

Clark, Richard E. (1985). Evidence for confounding in Computer-Based Instruction Studies: Analyzing the Meta-Analyses, Educational Communication and Technology Journal, 33, (4), 249-262.

Clark, Richard E. (1983). Reconsidering Research on Learning from Media, Review of Educational Research, 53, 4, 445-459.

Cook, Thomas, D. & Campbell, Donald T. (1979). Quasi-Experimental Design and Analysis Issues for Field Settings. Boston, MA: Houghton Mifflin Company.

Covey, Preston K. (1990). Art or Forgery? The Strange Case of Han Van Meegeren A Videodisc for Aesthetics and Art History. Journal of Computing in Higher Education, 2, (1), 3-31.

Dick, Walter, & Carey, Lou (1990). The Systematic Design of Instruction (3rd ed.). New York: Harper Collins Publishers.

Egan, Dennis E., Remde, Louis M., Gomez, Landauer, Thomas K., Eberhardt, Jennifer, & Lochbaum, Carol C. (1989). Formative Design-Evaluation of SuperBook, ACM Transactions on Information Systems, 7, (1), 30-57.

Gagné, Robert M., & Briggs, Leslie, J. (1979). Principles of Instructional Design (2nd ed.). New York: Holt, Rinehart & Winston.

Gay, Geri, & Raffensperger (1989). Considerations and Strategies in the Design of Interactive Multimedia Programs. Academic Computing, September, 24-58.

Gay, Geri, Trumbull, Deborah, Mazur, Joan (1991). Designing and Testing Navigational Strategies and Guidance Tools for a Hypermedia Program. Journal of Educational Computing Research,7, (2) 189-202.

Goldman, Elizabeth, & Barron, Linda (1990). Using Hypermedia to Improve the Preparation of Elementary Teachers. Journal of Teacher Education, 41, (3), 21-31.

Guba, Egon G. & Lincoln, Yvonna S. (1984). Epistemological and Methodological Bases of Naturalistic Inquiry. In G. F. Madaus, M. Scriven, D.L. Stufflebeam (Eds.), Evaluation Models Viewpoints on Educational and Human Services Evaluation, (pp. 311-333). Hingham, MA:Kluwer-Nijhoff Publishing.

Hammond, Nick (1989). Hypermedia and Learning: Who Guides Whom? In H. Maurer (Ed.), Computer Assisted Learning 2nd International Conference, ICCAL '89 Dallas Tx., USA May 1989 Proceedings, (pp. 167-181), Berlin, Germany: Springer-Verlag.

Hammond, Nick, & Allinson, Lesley (1988). Development and Evaluation of a CAL System for Non-Formal Domains: The Hitch-Hiker's Guide to Cognition, Computers in Education, 12, (1), 215-220.

Harvey, F. A., (1990). Integrating hypermedia and knowledge-based system technologies for training and information dissemination. In Proceedings of the 32nd international conference of the Association for the Development of Computer-Based Instructional Systems, San Diego, CA, October 28 -- November 1, 1990, 39. Columbus, OH: Association for the Development of Computer-Based Instructional Systems, Ohio State University.

Harvey, F. A., (1990). Designing hypermedia training materials to optimize learning. In Proceedings of the 32nd international conference of the Association for the Development of Computer-Based Instructional Systems, San Diego, CA, October 28 -- November 1, 1990, 57. Columbus, OH: Association for the Development of Computer-Based Instructional Systems, Ohio State University.

Harvey, F. A., (1989). Planning and developing an intelligent information integration and management system. In Proceedings of the 31st international conference of the Association for the Development of Computer-Based Instructional Systems, Arlington, VA, November 12-16, 1989, 267-71. Bellingham, WA: Association for the Development of Computer-Based Instructional Systems, Western Washington University.

Harvey, F. A., and Story, W.B., Jr. (1989). Developing hypermedia training materials for a national human services coordination program for at-risk youth. In Proceedings of the 31st international conference of the Association for the Development of Computer-Based Instructional Systems, Arlington, VA, November 12-16, 1989, 38-43. Bellingham, WA: Association for the Development of Computer-Based Instructional Systems, Western Washington University.

Higgins, Kyle, & Boone, Randall (1990), Hypertext Computer Study Guides and the Social Studies Achievement of Students with Learning Disabilities, Remedial Students, and Regular Education Students, Journal of Learning Disabilities, 23, (9), 529-540.

Horton, Steven V., Boone, Randall A., Lovitt, Thomas C. (1990). Teaching Social Studies to Learning Disabled High School Students: Effects of a Hypertext Study Guide, British Journal of Educational Technology, 21, (2), 118-131.

Jacob, Evelyn (1987). Qualitative Research Traditions: A Review, Review of Educational Research, 57, (1), 1-50.

Jones, Robert A. (1990). To "Criss-Cross in Every Direction"; or, Why Hypermedia Works, Academic Computing, January, 20-21;30;53.

Kerlinger, F. N. (1973). Foundations of Behavioral Research (2nd ed.). New York: Holt, Rinehart and Winston.

Kinzie, Mable B., & Berdel, Richard L. (1990). Design and Use of Hypermedia Systems, Educational Technology Research and Development, 38, (3), 61-68.

Kozma, Robert B. (1991). Learning with Media. Review of Educational Research, 61, 2, 179-211.

Landow, George P. (1989). Hypertext in Literary Education, Criticism, and Scholarship. Computers and Humanities, 23, 173-198.

Lanza, Antonietta, & Roselli, Teresa (1991). Effects of the Hypertextual Approach versus the Structured Approach on Students' Achievement. Journal of Computer-Based Instruction, 18, (2), 48-50.

Locatis, Craig, Letourneau, Gary, & Banvard, Richard (1989). Hypermedia and Instruction, Educational Technology Research and Development, 37, (4), 65-77.

Michalerya, W.D., Wilson, J.L., & Chen, S.S. (1990). "The Bridge Fatigue Investigator - A Knowledge-Based System for Bridge Inspectors," Proceedings of the Seventh Annual International Bridge Conference, Pittsburgh, PA, June, Paper No. 90-49.

Marchionini, Gary (1988). Hypermedia and Learning: Freedom and Chaos, Educational Technology, November, 8-12.

Martin, C. Dianne (1991). New Findings from Qualitative Data Using Hypermedia: Microcomputers, Control and Equity, Computers in Education, 16, (3), 219-227.

Mays, J. Terence, Kibby, Michael R., & Watson, Hugh (1988). The Development and Evaluation of a Learning-by-Browsing System on the Macintosh, Computers in Education, 12, (1), 221-229.

Merrill, M. D. (1983). Component Display Theory. In C.M. Reigeluth (Ed.) Instructional Design Theories and Models: An Overview of Their Current Status. New Jersey: Erlbaum.

Moline, Judi, Benigni, D., & Baronas, Jean (Eds.). (1990). Proceedings of the Hypertext Standardization Workshop January 16-18, 1990 National Institute of Standards and Technology. Washington, D.C: US Government Printing Office.

Reeves, Thomas C. (1991). Ten Commandments for the Evaluation of Interactive MultiMedia in Higher Education, Journal of Computing in Higher Education, 2, (2), 84-113.

Romiszowski, Alexander J. (1990). The Hypertext/Hypermedia Solution--But What Exactly is the Problem? In D.H. Jonassen & H. Mandl (Eds.). Designing Hypermedia for Learning, Berlin, Germany: Springer-Verlag.

Rossi, Peter, H., & Freeman, Howard E. (1989). Evaluations A Systematic Approach (4th ed.). Newbury Park: Ca: Sage Publications.

Saracevic, Tefko, Mokros, Hartmut, & Su, Louise. (1990). Nature of Interaction Between Users and Intermediaries in Online Searching: A Qualitative Analysis, In D. Henderson (Ed.), Proceedings of the 53rd Annual Meeting of the American Society for Information Science, (pp. 47-54). Medford, New Jersey: Learned Information, Inc.

Skinner, B.F. (1986). Programmed Instruction Revisited. Phi Delta Kappan 68, 103-110.

BEST COPY AVAILABLE

Spokane, A., Davison, R. M., Wise, D., and Hart, R. C. (1991). Formative evaluation of Project Operations trainer training and trainee training. (Tech. Rep. No. CIS/HDP 91.05) Bethlehem, PA: Lehigh University, Cities in Schools Hypermedia Development Project.

Story, William B., Jr., and Harvey, F. A., (1990). Techniques and procedures for the transparent automatic collection of data for hypermedia-based research. In Proceedings of the 32nd international conference of the Association for the Development of Computer-Based Instructional Systems, San Diego, CA, October 28 -- November 1, 1990, 393-5. Columbus, OH: Association for the Development of Computer-Based Instructional Systems, Ohio State University.

Stufflebean, Daniel L. (1984). The CIPP Model for Program Evaluation. In G. F. Madaus, M. Scriven, D.L. Stufflebeam (Eds.), Evaluation Models Viewpoints on Educational and Human Services Evaluation, (pp. 311-333). Hingham, MA:Kluwer-Nijhoff Publishing.

Tripp, Steven D., & Roby, Warren. (1990). Orientation and Disorientation in a Hypertext Lexicon, Journal of Computer-Based Instruction, 17, (4), 120-124.

Winn, William (1990). Some implications of cognitive theory for instructional design. Instructional Science, 19, 53-69.

Footnotes

[1]Guba and Lincoln (1984) use the term "rationalistic" to reference forms of conventional inquiry such as experimental and quasi-experimental forms of evaluation.

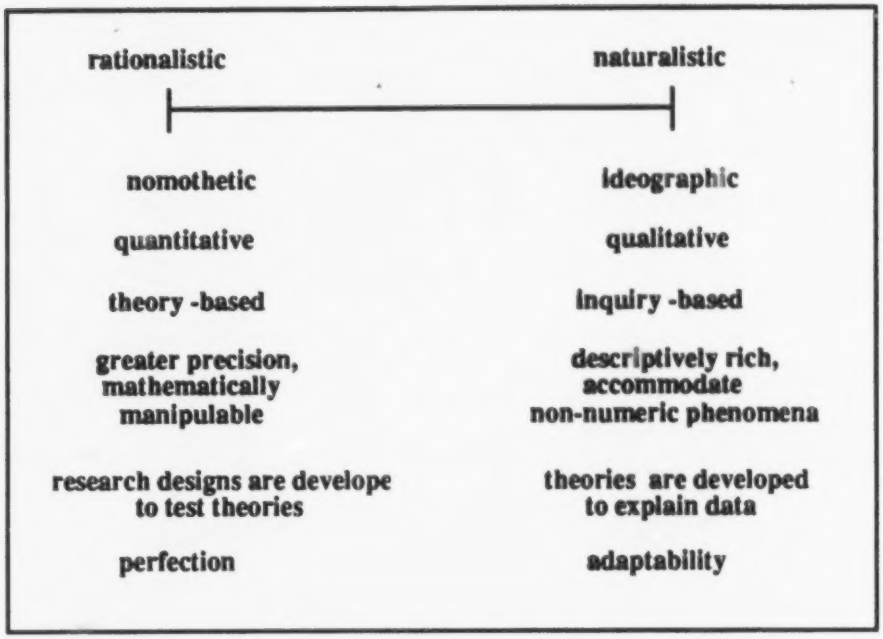

Figure 1: Comparison of rationalistic and naturalistic methodologies

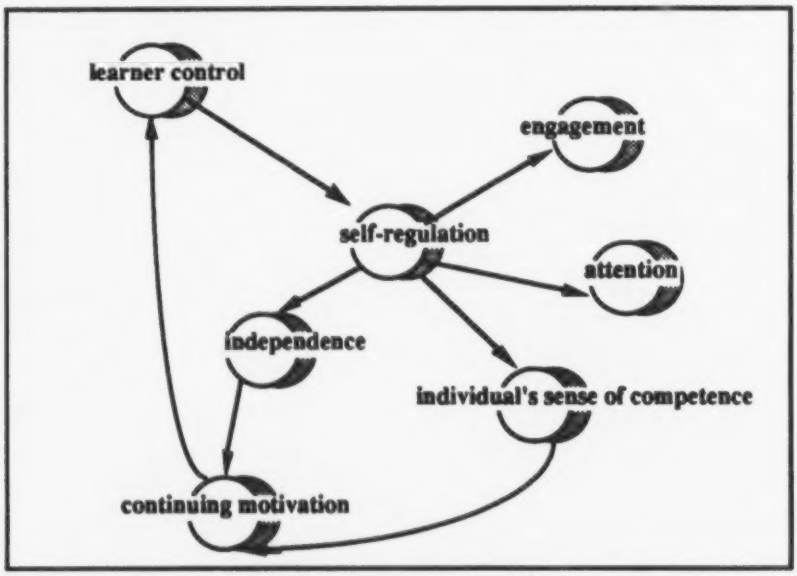

Figure 2: Relationship between learner control (primary characteristic) and self-regulation (secondary characteristic).

Table 1

Characteristics of hypermedia (Process) and expected Outcome (Product)

Cited Literature	Process Characteristic	Outcome Characteristic
Jones (1990)	Hypermedia more accurately reflects the way the mind actually works, i.e., thru the association of ideas rather than in a linear or hierarchical fashion.	Evidence supporting nonlinear over linear acquisition.
Kozma (1991)	Hypermedia facilitates application and transfer of complex knowledge to new situations via multiple exploration on either thematic or conceptual dimensions.	Evidence supporting knowledge representation along multiple instead of singular conceptual dimensions. Assembly of representations instead of retrieval of schema.
Kozma (1991)	Hypermedia provides learner with extension and construction ability.	Evidence supporting Hypermedia prompts learners to think about ideas, how they are interrelated, structured.
Locatis, Letourneau, & Banvard (1989)	Hypermedia requires linking with specific forms of instructional support.	Evidence which suggests success based on different forms of linking.
Locatis, Letourneau, & Banvard (1989)	Hypermedia node size and representation may affect use.	Evidence which indicates that different arrangements aid/impact success rates.
Locatis, Letourneau, & Banvard (1989); Hammond (1989)	Hypermedia open-ended experience may be different depending on previous exposure to content domain and aptitude level.	Evidence which indicates novice, intermediate, and expert skill require/use different treatment.
Marchionini (1988); Kinzie & Berdel (1990); Hammond (1989)	Hypermedia facilitates users' ability to control their own learning.	Evidence which indicates this is beneficial.
Marchionini (1988)	Hypermedia offers new ways of learning how to learn.	Evidence of increase in strategies for the learner.
Marchionini (1988)	Potential to alter roles of teachers and learners and facilitate interaction between them.	Evidence of Human-Computer-Intermediary support.

BEST COPY AVAILABLE

Table 2
Selected Evaluation Research Table 1 of 2

Study	Approach	Study Focus	Study Supported Research Hypothesis	Content	Subjects	Multiple Measures
Covey (1990)	Formative	Exploring knowledge evidentiary issues: attribution and critical analysis in study of visual art	no	Fine Art	College undergraduates	no
Egan et al. (1989)	Formative	Print search compared to computer search	no	Texas Legal Code	College undergraduates (with 1 or more statistics courses)	yes
Egan et al. (1989) Evaluation Study II.	Formative (between subjects measurement)	Added open ended questions, affective rating measures	yes	Texas Legal Code	College undergraduates (with 1 or more statistics courses)	yes
Egan et al. (1989) Evaluation Study III.	Formative	Smaller screen version compared to print search	yes	Texas Legal Code	College undergraduates (with 1 or more statistics courses)	yes
Gay, Trumbull, Mazur (1991)	Quasi-experimental (case study using volunteers)	Interaction of choice of search mode and Learning Study Skills	not clear	Cultural Entomology	College undergraduates	no
Goldman & Barron (1990)	Naturalistic	Affective measures, traditional test performance, teaching methods improvement	mixed	Mathematics teaching methodology	Preservice college students	no
Hammond, & Allinson (1988)	Formative	Performance	n/a	Cognitive Psychology	College undergraduates	no
Higgins & Boone (1990)	Quasi-experimental (pretest-posttest & retention, random assignment)	Comparing delivery modes: Lecture, Computer Study Guide, & both	no	Social Studies	High school students: regular, learning disabled, & remedial.	yes

Table 2a
Selected Evaluation Research Table 2 of 2

Study	Approach	Study Focus	Study Supported Research Hypothesis	Content	Subjects	Multiple Measures
Higgins & Boone (1990) Study II	Quasi-experimental (follow-up A-B-A)	Increase in test performance using study guide as follow-up instruction.	yes	Social Studies	Learning disabled, & remedial	yes
Horton, Boone, Lovitt (1990)	Quasi-experimental (small case)	Increase in comprehension using instructional sequencing in study guide	yes	Washington State History	Learning disabled high school students	yes
Landow (1989)	Naturalistic, ethnographic	Increase student's analytical skills, and develop ability to assimilate breadth of information	yes	Survey of English Literature, 1700 to the Present	College under-graduates	n/a
Lanza & Roselli (1991)	Quasi-experimental (post-test case study)	Effectiveness of hypertext instruction vs CAI	not clear	Pascal programming	College under-graduates	no
Mays, Kibby, & Watson (1988)	Experimental	Comparison of hypermedia and traditional teaching	no	Geology	College under-graduates	no
Mays, Kibby, & Watson (1988) Study II	Naturalistic	Protocol Analysis	not clear	Geology	College under-graduates	
Tripp & Roby (1990)	Experimental (post-test)	Visual metaphors increase learning, rote learning is not more effective with hypermedia	no	Japaneese-English lexicon	College graduates, and under-graduates	no

n/a = not applicable

Table 3
Criteria for Judging the Trustworthiness of the Findings of an Evaluation Study

Criteria for Judgement	Approach			
	Experimental	Quasi-Experimental	Formative	Naturalistic
Truth Value	Internal Validity	Internal Validity	Convincing	Credibility
Applicability	External Validity	External Validity		Transferability
Consistency	Reliability	Reliability		Dependability
Neutrality	Objectivity	Objectivity		Confirmability

Note. The data in column 1,2 and 5 are from "Epistemological and Methodological Bases of Naturalistic Inquiry" by E. G. Guba and Y.S. Lincoln , Evaluation Models (p. 326) in G.F. Madaus, M. Scriven, and D.L. Stufflebeam, (Eds.), 1983, Boston: Kluwer-Nijhoff Publishing. Adapted by the authors.